INFORMATIONAL LEAFLET NO. 230

ACCURACY OF SCALE PATTERN ANALYSIS IN SEPARATING MAJOR STOCKS OF SOCKEYE SALMON (Oncorhynchus nerka) FROM SOUTHERN SOUTHEASTERN ALASKA AND NORTHERN BRITISH COLUMBIA

Ву

Scott L. Marshall Glen T. Oliver David R. Bernard

and

Scott A. McPherson

STATE OF ALASKA

Bill Sheffield, Governor

DEPARTMENT OF FISH AND GAME

Don W. Collinsworth, Commissioner

P. O. Box 3-2000, Juneau 99802



ACCURACY OF SCALE PATTERN ANALYSIS IN SEPARATING MAJOR STOCKS OF SOCKEYE SALMON (Oncorhynchus nerka) FROM SOUTHERN SOUTHEASTERN ALASKA AND NORTHERN BRITISH COLUMBIA

Ву

Scott L. Marshall Glen T. Oliver David R. Bernard

and

Scott A. McPherson

Alaska Department of Fish and Game Division of Commercial Fisheries Statewide Stock Biology Group

January 1984

TABLE OF CONTENTS

<u>Pag</u>	Je
LIST OF FIGURES	İ
LIST OF TABLES	i
ABSTRACTiii	i
INTRODUCTION	ĺ
METHODS	2
Overview	2
Model Development and Use	2
Trottimition y companies to the transfer of th	4
Historical Models and In-season Estimation of Interception	1
	† 7
Scale Collection and Measurement	2
Scale Correction and measurement	J
Computations	3
RESULTS	ļ
Differences in Scale Patterns	1
Preliminary Comparisons	1
Historical Models: Nation of Origin Models (NOM) 1	İ
Historical Models: Stikine Models (SM)	5
Scale Selection)
DISCUSSION)
Differences in Scale Patterns)
Other Stocks	5
Appropriate Models	
RECOMMENDATIONS	,
ACKNOWLEDGMENTS	5
LITERATURE CITED	7

LIST OF FIGURES

Figure		Page
1.	Fishery management districts in southern Southeastern Alaska and northern British Columbia	3
2.	Rivers in southern Southeastern Alaska and northern British Columbia that have major populations of sockeye salmon	5
3.	Numbers and kinds of models used to measure the accuracy of historical models and the effect on accuracy of rules by which scales are selected for model construction	6
4.	Typical scales for age 1.3 and 2.3 sockeye salmon showing the zones used to measure scale patterns	9
5.	Typical scales from age 1.3 sockeye salmon from Alaska (right) and from Canada (left)	
6.	Typical scales from age 2.3 sockeye salmon from Alaska (right) and from Canada (left)	
7.	Differences for the most significant variables in the age-specifi NOM between sockeye salmon in Alaskan rivers and in Canadian rivers	
8.	Comparative lengths of rivers in southern Southeastern Alaska and northern British Columbia that have major populations of sockeye salmon	

LIST OF TABLES

Table		Page
1.	Preliminary comparisons of scale patterns of sockeye salmon from 24 Alaskan rivers and the Nass, Skeena, and Stikine Rivers	15
2.	The mean accuracy (%) attained and the variables selected for age-specific and age-pooled Nation of Origin Models (NOM) for sockeye salmon	17
3.	The mean accuracy (%) attained and the variables selected from age-specific and age-pooled Stikine Models (SM) for sockeye salmon	18
4.	Number of scales correctly and incorrectly classified by SM for age 1.2, 1.3, and 2.2 sockeye salmon	19
5.	Comparison of mean accuracies (%) attained from models construct from scales selected from stocks of sockeye salmon from both Alaska and from Canada according to four rules	
6.	Comparison of four fules for selecting scales from stocks of sockeye salmon from Alaska and Canada against sets of scales with known origins	22

ABSTRACT

The feasibility of using scale patterns and linear discriminant functions to estimate the contribution of Alaskan and Canadian stocks of sockeye salmon (Oncorhynchus nerka Walbaum) to the fisheries of southern Southeastern Alaska and northern British Columbia was examined using data collected in 1982. nificant and persistent differences were found in the patterns of scale growth during freshwater and early marine life history between stocks originating in Alaska and Canada. Sockeye salmon from Alaskan stocks grew less and slower during their lacustrine residence than did fish from Canadian stocks. Also, Alaskan fish rarely exhibited spring plus growth, while Canadian fish almost always did. Because jackknife accuracy for the linear discriminant function remained high when data for each of four age classes was pooled (93.3% average for four age-specific models vs 88.2% for one age-pooled model), the variation in patterns between nations is greater than within nations across years. When the Stikine River stocks were separated from other Canadian stocks, a small penalty in accuracy was paid; scales from Stikine River fish were misclassed most often as being from other Canadian stocks. A sensitivity analysis on the way in which scales are selected for model construction was conducted with a second set of scales of known origin. Because model accuracy proved robust to the manner of scale selection, no prior knowledge regarding migratory pathways, stock abundance, or age composition was required to draw samples to represent each nation. We conclude that scale pattern analysis is a cost effective and efficient method for estimating the contribution of each nation's stocks to the fisheries of the other. Because interannual variability in scale patterns is small, programs with historical models can be developed to provide estimates of interception rates during each fishery.

KEY WORDS: sockeye salmon stock identification, interception rates, scale patterns, salmon stock classification accuracy, scale sampling design, identification of Canadian and Alaskan sockeye stocks

INTRODUCTION

During the last two decades, the governments of the United States of America and of Canada have been negotiating a treaty to cooperate in the management, research, and enhancement of Pacific salmon (Oncorhynchus sp.) stocks that are harvested by nationals of both countries. Interception of salmon bound for one country's rivers as they migrate through the territorial waters of the other country is one of the negotiated issues. In southern Southeastern Alaska, limiting the magnitude of interception of sockeye salmon (O. nerka) bound for the Nass, Skeena, and Stikine Rivers is of concern to Canada. If fishery managers in Alaska are to answer this concern while permitting harvest of suplus production from Alaskan rivers, they require knowledge of the temporal and spatial variation in interception rates of sockeye salmon from the Nass, Skeena, and Stikine Rivers in Alaska waters. A similar situation exists with respect to interceptions of Alaskan sockeye salmon by Canadians.

In 1982, an international research program was started to assess the feasibility of several methods of estimating the numbers of salmon bound for rivers in one country but intercepted by fisheries of the other. Each tested method uses characteristics of captured fish to indicate where those fish would have gone had they not been caught. For instance, if the parasite <code>myxobolus neurobius</code> is found only in the brains of fish bound for rivers in Alaska, a simple allocation can be made according to the incidence of this parasite in samples of brain tissue taken from captured salmon. Besides the incidence of pathogens, differences in genotypes, and scale patterns have been proposed as attributes upon which methods to allocate catches can be based. Also, man-made attributes, such as tags, have been proposed as the basis for methods to estimate interception rates.

Scale pattern analysis is based on differences in the arrangement of circuli on scales. Because the pattern of circuli is a history of the growth of a fish, salmon with the same history have similar scale patterns, while those with different histories have different patterns. The greater the difference among scale patterns, the better one pattern can be distinguished from another. scale pattern analysis, subsamples of scales, one subsample from each escapement sample, are combined to represent the groups of stocks to be separated in the catch. Patterns of circuli are measured in several ways on each subsampled scale, thereby building data sets of variables that describe the typical pattern for each group of stocks. Next, these data sets are compared against each other with statistical techniques to build a decision rule (discriminant function) which can be used to classify scales as belonging to one of the groups. the discriminant function is tested with scales from fish of known origin, and correction factors are calculated from the errors made with the discriminant function. Finally, the discriminant function and its correction factors are used to classify fish of unknown origin from the catch to one of the groups. The incidence of fish classified to a run is the interception rate for that group.

In this report we answer the following questions: Can scale pattern analysis be used to distinguish fish bound for Alaskan versus Canadian (primarily the Nass, Skeena, and Stikine) rivers? First, are the scale patterns on fish with origins in Alaskan territory different than the patterns on scales from fish from Canadian

territory? If differences in patterns exist, scale pattern analysis is possible. Next, are differences greater between nations or within nations? If differences are greater between nations, then scale pattern analysis can provide an accurate decision rules that parallel management concerns. And finally, if differences in scale patterns do exist, are these differences persistent from year to year? If differences are persistent, scale pattern analysis based on historical models could be used during a fishing season to calculate interception rates as the season progresses.

METHODS

Overview

We considered two, commonly used methods of multivariate discriminant analysis: the parametric linear discriminant analysis (Fisher 1936) and the nonparametric nearest-neighbor analysis (Lachenbruch 1975). Because variables used in the analyses proved normally or near-normally distributed and because preliminary analysis with both techniques showed linear discriminant analysis had the higher accuracy (only marginally so) with our data, the more computationally complex and costly nearest-neighbor analysis was not used. Oliver et al. (1983) and the RESULTS section of this report contain more detailed ussport for linear discriminant analysis as the appropriate technique for the 1982 data.

The accuracy of decision rules generated from linear discriminant analysis on scale patterns is an indication of the feasibility of this technique to correctly identify the origin of sockeye salmon harvested in commercial fisheries. One kind of accuracy is estimated during the construction of the linear discriminant function, hereafter called the model, for each group individually and for all groups as a mean accuracy. Once minimum sample sizes are met, this first kind of accuracy is the principal component of the variance of estimated interception rates (Pella and Robertson 1979). A second kind of accuracy is the ability of the model to correctly classify the origins of a set of scales not used in model construction. Both kinds of accuracy were investigated for this report.

Model Development and Use

Discussions with research and management biologists from both Canada and Alaska exposed the need for two comprehensive models, a National Origin Model (NOM) and a Stikine Model (SM). The NOM is useful in those fisheries where the mixed stocks are only from rivers that lie totally within the boundaries of one or the other nation. Preliminary results of an international tagging program conducted in southern Southeastern Alaska in 1982 and previous tagging studies (Rich and Morton 1930; Verhoeven 1952; Noerenberg 1959; Logan 1967; Simpson 1968) indicate that a NOM could be useful for fisheries in regulatory Districts 101 through 104 and 107 in Alaska and in Districts 1, 3, 4, and some of the northern subdistricts of District 31 in Canada; all are districts that contain few or no fish headed for the Stikine River (Figure 1). Because estimation of the catches of sockeye salmon bound for the Stikine River, a transboundary river, is a separate and an important issue, the Stikine Model must do all that the NOM can do plus separate sockeye salmon heading to the Stikine River from the remaining Canadian stocks. The SM

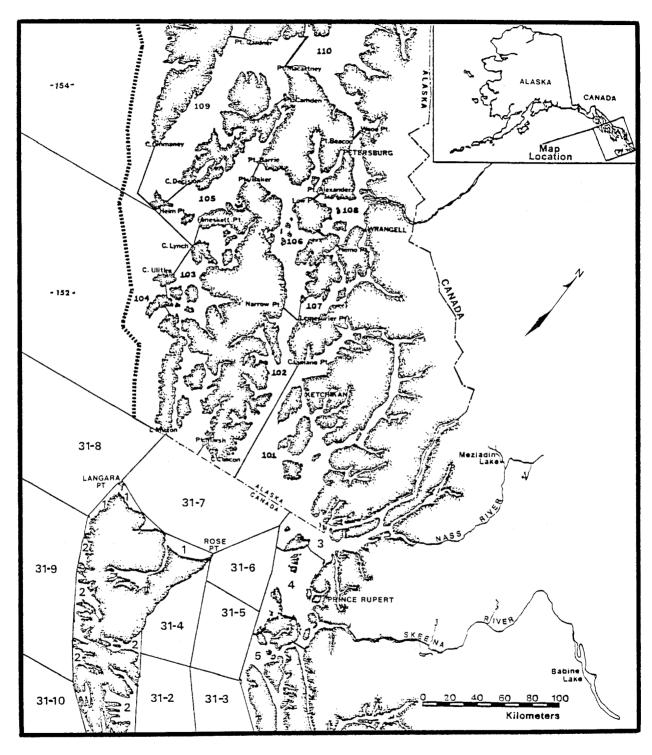


Figure 1. Fishery management districts in southern Southeastern Alaska and northern British Columbia.

could be most useful in fisheries in Alaskan Districts 105, 106, and 108 where fish from the Stikine River have been most commonly found, but could be used in all Alaskan and Canadian districts where fish from the Stikine River are present or not.

Scales for the models were taken from 28 Alaskan rivers with the largest 1982 escapements of sockeye salmon between Frederick Sound and Dixon Entrance (McGregor 1983) and from the Nass, Skeena, and Stikine Rivers (Figure 2). For the NOM, Alaska stocks are those from rivers that are toally within US territory, and Canadian stocks are those from the Nass, Skeena, and Stikine Rivers. In the SM, the stocks of the Stikine River are a separate group. Although preliminary results from international tagging studies done in 1982 show that some fish from stocks from as far north as Yakutat, Alaska and from as far south as the Fraser River, British Columbia were caught in outer coastal fisheries, the incidence of these fish is negligible, and they were not considered for this analysis.

Preliminary Comparisons:

One linear discriminant model for age 1.3¹ sockeye salmon was built to show which populations of Alaska fish had scales most like those from Canada and vice versa. The Stikine, Nass, Skeena, and each Alaskan river were each considered a separate origin in the model. Ninety-six scales from the Stikine River, 110 from the Nass River, and 101 from the Skeena Rivers and about 50 from each of 24 Alaska rivers² were randomly selected from samples.

Historical Models and In-season Estimation of Interception Rates:

Because annual variation in scale patterns affects the ability of historical models to distinguish scales from the two nations, a series of NOM and SM were developed to test the significance of these effects on model accuracy (Figure 3). First, scales were chosen from Alaskan and Canadian rivers for these models according to the Equal Probability Rule as described in the following section of this report. Next, NOM and SM were developed for each major age class in these stocks which represented 3 year classes (1976, 1977, and 1978). Because ages 1.2, 1.3, 2.2, and 2.3 represented 99.7% of the 1982 catch (McGregor 1983), these were the only age classes used in the analysis. Next, scales were pooled by similar freshwater ages and then over all ages in the analysis. The decrease in accuracy from age-class specific models to age-pooled models is an estimate of the effect of annual variation in scale patterns on the accuracy of the models and subsequently on the usefulness of historical models for in-season estimation of interception rates.

¹ European Formula - Numerals preceding the decimal refer to the number of freshwater annuli; numerals following the decimal are the numbers of marine annuli. Total age is the sum of these two numbers plus one.

Not all 28 Alaskan runs have large numbers of age 1.3 sockeye salmon; runs to Leask, Kushneahin, Shipley, and Kah Sheets Lakes do not. Also, only 47 samples are available from Helm Lake.

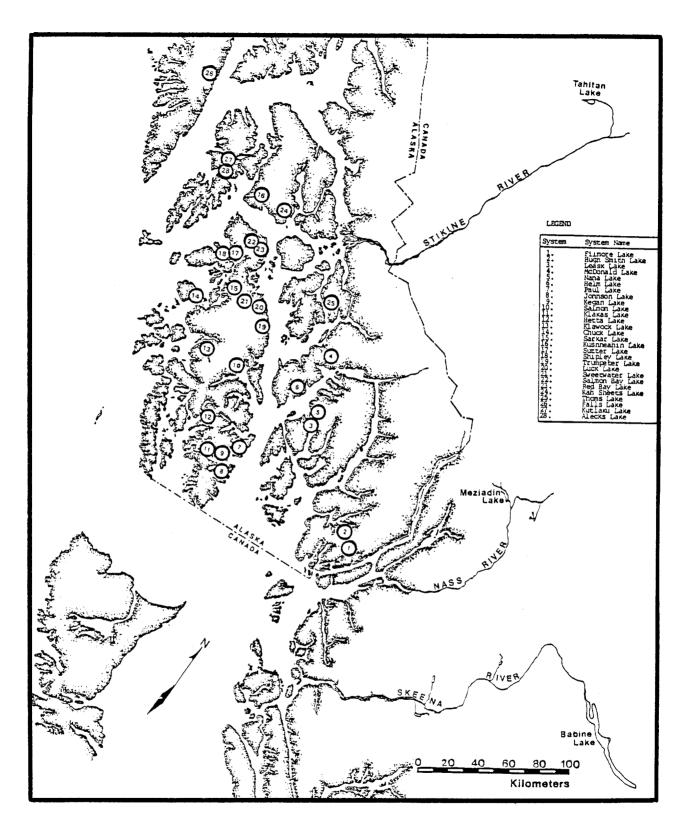


Figure 2. Rivers in southern Southeastern Alaska and northern British Columbia that have major populations of sockeye salmon.

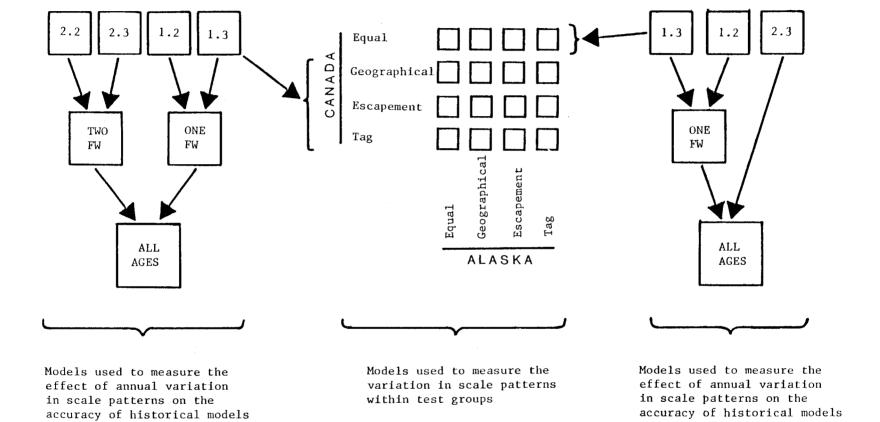


Figure 3. Numbers and kinds of models used to measure the accuracy of historical models and the effect on accuracy of rules by which scales are selected for model construction. Each box corresponds to a model. Ages are listed inside the boxes for age-specific and age-pooled models; ONE FW or TWO FW refers to freshwater ages. See text for explanations of rules for scale selection.

-6-

Scale Selection and Model Accuracy:

If variation in scale patterns within territorial regions is large relative to that among regions, scales with which to build models to discriminate among-region differences in patterns must be selected according to the sizes of stocks, their migration patterns, and possibly their age compositions to provide a representative sample of scales for the region. Conversely, if within-region variation in scale patterns is small relative to among-region variation, scales can be chosen for the models without regard to these other attributes of the stocks without significantly affecting the accuracy of models.

We built several NOM and SM based on scales that were selected with different rules to measure the effect of scale selection on model accuracy, and tested their accuracy on other scales of known origin. Four rules for selecting scales were used to build 16 models (Figure 3), and each model was used to separate three groups of test scales taken from escapements in the following proportions 90% Alaska 10% Canada; 50% Alaska 50% Canada; and 10% Alaska 90% Canada. Only scales from fish aged 1.3 were used in building models and in test populations. The test scales are substitutes for scales from catches made in Districts 101 and 102 and were selected from those escapements within Alaska and Canada that contributed to fisheries in these districts according to preliminary results from the 1982 tagging program. Because the tagging program found no fish from the Stikine River in these districts in 1982, only scales from the Nass and Skeena Rivers were used as test scales from Canada. The four rules are as follows below:

- 1) Equal Probability. Scales were chosen at random from all rivers in the study area. To select N scales from the 24 Alaskan runs used in a model, we chose N/24 scales at random from each of the 24 runs, and to select N scales from the three Canadian rivers, randomly chose N/3 from each. Note that under this rule, scales from the Stikine River were used to build the SM even though no scales from the Stikine River are in the test population. Also note that unless otherwise stated, SM and NOM in the test of the historical models are built with scales selected by the Equal Probability Rule.
- 2) Geographical Probability. Scales were chosen at random from all important runs that contributed to the fisheries in Alaskan Districts 101 and 102 according to tagging studies conducted in 1982. To select N scales from the six major and four minor Alaska runs¹ that have fish caught in the districts, we chose N/7 scales randomly from each of the six major rivers and the one group of minor streams. To select scales from Canadian rivers by this rule, we ignored all scales from the Stikine River and selected equal numbers (N/2) from the Nass and the Skeena Rivers.
- 3) Escapement Probability. Scales were chosen according to the Geographical Probability Rule, only the number of scales chosen from each run is a

Major runs to Hugh Smith, McDonald, Naha, Helm, Kegan, and Karta Lakes; minor runs to Filmore, Johnson Cove, and Paul Lakes.

product of the size of its escapement of age 1.3 fish. To select N scales from M runs each with an escapement E, we randomly chose NE, \leq E from the ith river.

Tag Probability. Scales were chosen according to the Geographical Probability Rule, only the number of scales chosen from each run is a product of the contribution it made to the catch according to tagging studies conducted in 1982. To select N scales from M runs each with tags T caught in the fishery, we randomly chose NT_i ≤T from the ith river. Because no tags from Districts 101 and 102 were recovered from the escapement to one large Alaskan river, this run was excluded for this rule¹.

If models built with scales selected by the Equal Probability Rule have the same accuracy as models built with scales selected by the other rules, variation of scale patterns within groups in the comparison is small relative to variation among groups, and scales can be selected without regard to run strength or migratory patterns.

Scale Collection and Measurement

Samples were collected over the entire season at weirs or on the spawning grounds in Alaska, from the gillnet fishery in the lower Stikine River, and from the test fisheries in the Nass and Skeena Rivers. Scales were taken from the preferred area (see Clutter and Whitesel 1956), mounted on gum cards, and impressions made in cellulose acetate.

Scales were measured according to zones that parallel the age of the fish. Criteria used to determine age were essentially those of Mosher (1968). The first two zones correspond to the first 2 years spent in freshwater (note that there is no second zone on age 1.2 and 1.3 fish), the third zone corresponds to plus growth realized during the spring of smoltification, and the fourth zone corresponds to the first year in the sea (Figure 4). When the edges of resorted scales precluded direct interpretation of marine age, marine age was assigned with lengthfrequency histograms for individual escapements (see Tesch 1970). Although all readable scales were aged, detailed measurements were made on a randomly selected subset from each sample of scales. Scale impressions were projected onto a digitizing tablet at 100x magnification using equipment similar to that described by Bilton (1970) and modified by Ryan and Christie (1976). The size of the zones (variables named IDs), the number of circuli in the zones (variables named NCs), and the distance to circuli within zones (variables named TWOs, FOURs, SIXs, and EIGHTs) were measured along a line starting at the focus with an angle of 20 degrees from the long axis and perpendicular to the sculptured field.

Computations

Scale measurements were grouped according to age and location to produce mean vectors and variance-covariance matrices for the models. Scale measurements were put into a series of data matrices:

¹ The run to Helm Lake.

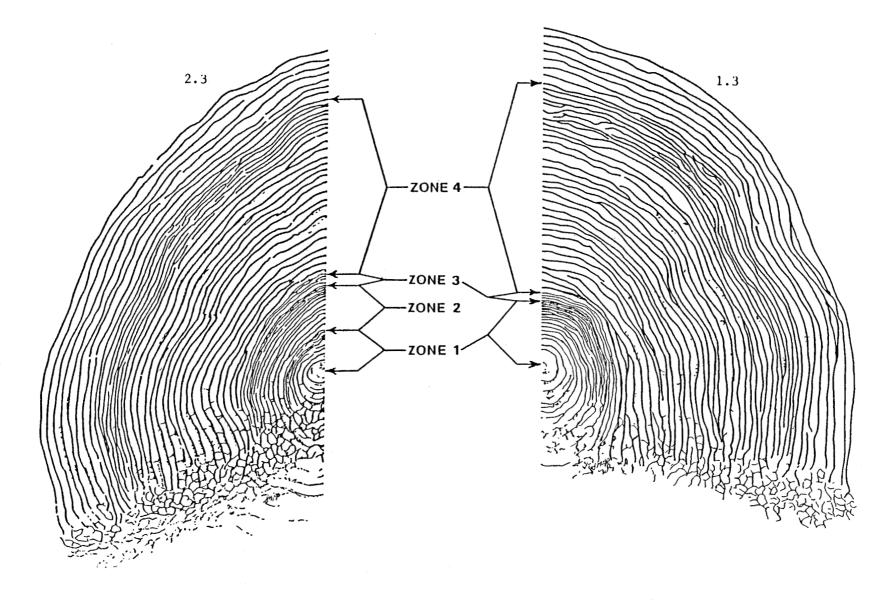


Figure 4. Typical scales for age 1.3 and 2.3 sockeye salmon showing the zones used to measure scale patterns.

where each row represents all measurements of (p) variables for one fish and each column represents measurements of one variable for all (n) fish. A matrix and a mean vector of variable measurements were generated for each combination of age class and location. For the age-specific NOM, eight data matrices were generated (1.2 Alaska, 1.2 Canada, 1.3 Alaska, 1.3 Canada, 2.2 Alaska, 2.2 Canada, 2.3 Alaska, and 2.3 Canada) to produce four models to separate Alaska from Canada for fish aged 1.2, 1.3, 2.2, and 2.3. For the age-specific SM, nine data matrices were generated (1.2 Alaska, 1.2 Stikine, 1.2 Nass-Skeena, 1.3 Alaska, 1.3 Stikine, 1.3 Nass-Skeena, 2.3 Alaska, 2.3 Stikine, and 2.3 Nass-Skeena); the Stikine River contains almost no fish age 2.2. Age-pooled NOM and SM had fewer data matrices, each a pooled combination of the data matrices for each of the constituent ages in the pooled model.

Linear discriminant functions were built and their accuracy tested with procedures available in the BioMedical Computer Programs P-series (Dixon and Brown 1979) on the Honeywell Computer at the University of Alaska. In a typical analysis with BMDP, mean vectors and variance-covariance matrices of measurements are derived from the data matrices. Variance-covariance matrices are pooled across the populations to be discriminated, inverted, and the differences among the mean vectors of these populations found. The discriminant function is built one variable at a time with the next most discriminating variable included next. For each stage of construction of the function, a jackknife procedure¹ is used to estimate the individual group and mean accuracy of the model and to build a correction matrix to compensate for its errors². The comparison of age-specific and age-pooled models is based on the mean accuracy from the jackknife procedure.

Application of a model and its correction matrix ${\bf C}$ to a set of scales from the catch provides estimates of interception rates (see Cook and Lord 1978). A sample of scales from a catch containing a mixture of groups can be represented by a

A discriminant function is built using all samples but one. The function is then used to estimate the origin of that one, omitted fish. Since the origin of the omitted sample is known, the accuracy of the function is tested. The procedure is then repeated only with a new sample excluded, and continues in this fashion until the origins of all samples are estimated and the accuracy of the model measured. Estimates of accuracy derived with the jackknife procedure are slightly biased, but negligibly so, as long as sample sizes are large (>25).

The correction matrix is a square matrix with one column and one row for each group. The element in the ith row jth column of the matrix is the fraction of scales in group j that were classified as being from group i through the jack-knife procedure. Diagonal elements in the matrix represent correctly classified scales, while off-diagonal elements represent misclassified scales.

vector \mathbf{u} whose elements are the proportions that each group actually represents in the catch. Use of the model to distinguish scales of unknown origin provides an estimate $\hat{\mathbf{u}}$ which is related to \mathbf{u} in the following manner:

$$u^{\dagger}C = \hat{u}^{\dagger}$$

Since C and \hat{u} are known:

$$u = c^{-1}\hat{u}$$

where **u** now contains the corrected estimates of the interception rates for each group. For this procedure, Pella and Robertson (1979) developed a means of calculating 90% confidence intervals for the estimated interception rates. Estimates of interception rates and their confidence intervals were used to compare accuracies of models based on scales selected according to different rules.

RESULTS

Differences in Scale Patterns

The differences between patterns on scales taken from sockeye salmon from each nation in 1982 are apparent to the naked eye (Figures 5-6). Scales from Alaska have small freshwater zones for 1 and for 2-year residents in freshwater. The corresponding freshwater zones on scales from Canadian rivers are larger. Also, zone three, the zone that corresponds to plus growth, is obvious in Canadian scales but is nonexistent on most Alaskan scales. Histograms of measurements (Figure 7) of the distances to circuli in the first freshwater and the plusgrowth zones (zones 1 and 3, respectively) on randomly selected scales from both countries show that the freshwater zones on scales from Canada are significantly larger than the same zones on scales from Alaska. Plus growth is about the same for sockeye salmon from Alaska and from Canada when it exists, but only 18% of the scales from Alaska have plus growth compared to 92% of scales from Canada (Figure 7).

Preliminary Comparisons

Few Alaskan rivers had scales that were misclassified as being from the Stikine, Nass, or Skeena Rivers, and few scales from these rivers were misclassified as being from Alaska (Table 1). Scales from the Skeena River misclassified as being from Alaska rivers (and vice versa) least often. Only 10% of scales from the Stikine River and 16% of those from the Nass misclassified as being from Alaska rivers. Of all the listed Alaska rivers in Table 1, only Filmore Lake and Helm Lake had a significant portion of scales misclassified as belonging to one of the large Canadian rivers. According to results from past tagging studies, sockeye salmon from Filmore and from Helm Lakes have contributed to the fisheries in District 101 and 102, especially the larger run to Helm Lake.

Historical Models; Nation of Origin Models (NOM)

The NOM, one for each of the four major age classes, correctly classified 87.8 to 92.8% of the sockeye salmon in the jackknifed sample to either Alaskan or

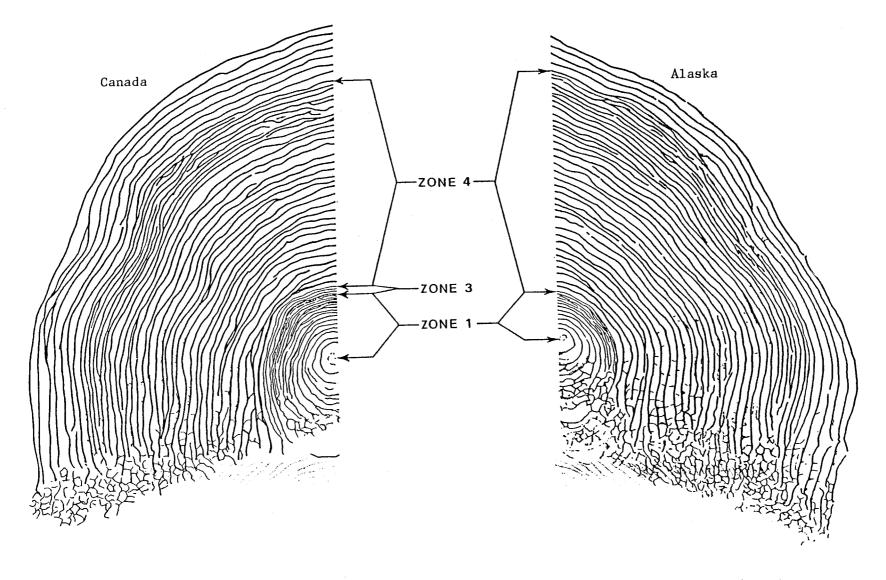


Figure 5. Typical scales from age 1.3 sockeye salmon from Alaska (right) and from Canada (left).

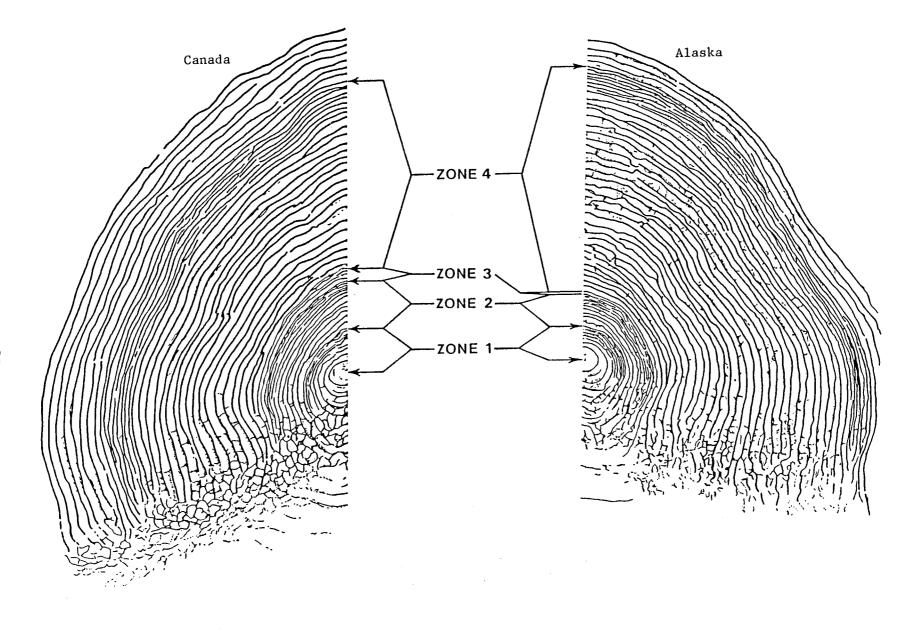


Figure 6. Typical scales from age 2.3 sockeye salmon from Alaska (right) and from Canada (left).

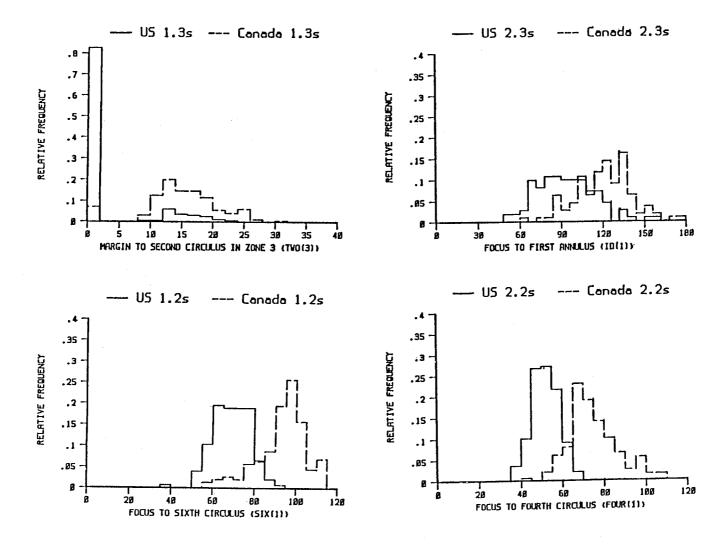


Figure 7. Differences for the most significant variables in the age-specific NOM between sockeye salmon in Alaskan rivers and in Canadian rivers.

Table 1. Preliminary comparisons of scale patterns of sockeye salmon from 25 Alaskan rivers and the Nass, Skeena, and Stikine Rivers. Arrows show the direction of misclassification. The number of rivers in the Others category follows in parentheses. The numbers on the arrows are the percent misclassified; for Others, these numbers are averages for the category. Those Alaskan rivers not specified or represented in Others are associated with no misclassifications.

Major Misclassificat	tion	Minor Misclassification			
Nass River <> 5%	Filmore Lake	1% Nass River <> Others (6) 11%			
2% Skeena River <> 4%	Thoms Lake Falls Lake	Skeena River $\stackrel{0\%}{}$ Others (0) 4%			
Stikine River>	Helm Lake Klakas Lake	3% <			

Canadian waters with only three variables included in the function (Table 2). Asymptotic accuracy of the models ranged from 88.7 to 96.0%, while the accuracy with only the most significant variable in a model (a univariate discriminant function) ranged from 75.2 to 89.8%. In contrast, classification by random chance alone would be correct only 50% of the time with two choices. Measurements from the zone corresponding to first-year lacustrine growth $\{SIX(1), FOUR(1), and ID(1)\}$ and of the zone corresponding to plus-growth $\{TWO(3)\}$ were the most important variables in each NOM (Figure 7). The differences in the variables indicated that sockeye salmon spawned in 1976 (2.3), 1977 (1.3 and 2.2), and 1978 (1.2) grew larger in Canadian lakes and had more plus growth there as well.

When variables were pooled over ages, differences in zones representing lacustrine growth over the years remained important to accuracy. Models for the same freshwater ages classified 88.4 to 89.3% of the sockeye salmon in the sample correctly with only two or three of the most significant variables included (Table 2). Asymptotic accuracy of these two-choice models ranged from 89.6 to 93.2%, while the accuracy with only the most significant variable ranged from 81.4 to 83.9%. As before with two choices, classifications by random chance alone would be correct 50% of the time. Measurements within the zone corresponding to first-year lacustrine growth {FOUR(1), SIX(1)} were the most important variables in the models. If age is totally ignored, the accuracy of the all-age model with three variables was 87.2%, its asymptotic accuracy was 88.2%, and its accuracy with only the most significant variable {FOUR(1)} was 79.7% (Table 2).

Historical Models: Stikine Models (SM)

The SM, one for each of the three major age classes 2 , correctly classified 76.7 to 80.5% of the sockeye salmon in the jackknifed sample to either Alaskan rivers or the Stikine River or to the Nass and Skeena Rivers with only three variables included in the models (Table 3). Asymptotic accuracy (with all ten variables) of the models ranged from 83.3 to 87.8%, while the accuracy with only the most significant variable in the models ranged from 58.4 to 70.8%. In contrast, classifications by random chance alone would be correct only 33% of the time with three choices. Measurements of the size of the zone corresponding to first $\{SIX(1)\}$ and $\{SIX(2)\}$ of lacustrine growth were the most important variables in the models. The differences in the variables indicated that sockeye salmon spawned in 1976 (2.3), 1977 (1.3), and 1978 (1.2) grew larger in Canadian lakes. For the three SM for individual age classes, scales from the Stikine River misclassified more often as being from the Nass or Skeena Rivers, and scales from the Nass or Skeena Rivers more often misclassified as being from the Stikine River (Table 4). Classification of scales from Alaska were highly accurate.

When variables were pooled over ages, the importance or differences in zones representing lacustrine growth remained. The SM for all pooled ages classified 72.8% of the sockeye salmon in the sample correctly with only three of the most significant variables (Table 3). Asymptotic accuracy of these three-choice models ranged

Asymptotic accuracy is attained when new additions of variables to the model do not perceptably change its accuracy.

Few sockeye salmon age 2.2 were found in the escapement to the Stikine River in 1982.

Table 2. The mean accuracy (%) attained and the variables selected from agespecific and age-pooled Nation of Origin Models (SM) for sockeye salmon. The more significant variables are those with the lower rank in the order of significant variables. Asymptotic accuracy is noted by arrows leading from values. Minor increases in accuracies are omitted.

			RANK	ORDER OF	e signi.	FICANT	VARIABLES			
AGE CLASS	1	2	3	4	5	6	7	8	9	10
1.2	89.1	91.9							92.5	>
1.3	86-0	88.3	89.8	90.5					91.8	 >
2.2	89.8	92.0	92.8							96.0
2.3	75.2	79.3	87.8				88.7 -	>		
1.2 & 1.3	81.4	86.8	89.3					89.6 -	 >	
2.2 & 2.3	83.9	88.4							93.2	
All	79.7	84.9	87.2							88.2
Variables i	n sequenc	ce added								
1.3 TWO (3) 2.2 FOUR(1	,ID(4),F0	OUR(1),NC(,NC(4),ID(3),NC(4), 4),NC(2),	NC(3), TW FOUR(3), TW TWO(4), ID(OUR(3), EIG	0(1),TWO(4 1),NC(1),I	EIGHT (1.	
1.2 & 1.3 2.2 & 2.3							HT (4)			
All FOUR(),NC(4),	IWO (3) , IWO	(4),ID(3)	,NC(3),EIG	HT(4),SIX	(1) ,NC(1)	,ID(1)	1		

Table 3. The mean accuracy (%) attained and the variables selected from agespecific and age-pooled Stikine Models (SM) for sockeye salmon. The more significant variables are those with the lower rank in the order of significant variables. Asymptotic accuracy is noted by arrows leading from values. Minor increases in accuracies are omitted.

			RA	NK ORDI	ER OF SI	GNIFICA	NT VARI	ABLES		
AGE CLASS	1	2	3	4	5	6	7	8	9	10
1.2	70.8	74.2	76.7							83.3
1.3	70.8	75.2	76.3							78.2
2.3	58.4	78.2	80.5							87.8
1.2 & 1.3	62.9	71.6	73.4							78.2
1.2 & 1.3 & 2.3	60.4	69.2	72.8			- 			77.3	>
Variables i	n sequenc	ce added				·····				
1.3 TWO(3)	,ID(4),FC	VO(2),ID(1) OUR(1),NC(3 ID(3),IWO(2	NC (4) ,FO	UR(3),TWO	(1),TWO(4)	EIGHT (4)	, ,			
1.2 & 1.3	FOUR(1),	C(2),ID(3)	,TWO(2),ID	(1) ,NC(1)	,TWO(1),ID	(2) ,TWO(3)	,EIGHT(3)			
1.2 & 1.3 &	2.3 FOU	JR(1),ID(3)	,TWO(2),TW	O(3),SIX(1) ,NC(3) ,E	GHT(3),NC	(1),ID(1)			

Table 4. Number of scales correctly and incorrectly classified by SM for age 1.2, 1.3, and 2.2 sockeye salmon. Each SM is based on a discriminant function with 10 variables. Classification accuracies were calculated with a jackknife procedure. Underlined numbers are correctly classified.

Actual Origin		ssified Or Nass - Skeer	rigin na/Stikine	Accuracy		
Alaska Nass-Skeena Stikine	91 5 9	6 89 8	<u>89</u> 7		10 variables Age 2.3	
Alaska Nass-Skeena	169 6	3 156	28 38	85 78	an Accuracy 10 variables	
Stikine ————————————————————————————————————		36	10	 92	Age 1.3 an Accuracy 10 variables	
Nass-Skeena Stikine	3 13	136 28	21 1 19	85 74	Age 1.2	

84% Mean Accuracy

from 77.3 to 78.2, while the accuracy with only the most significant variable ranged from 60.4 to 62.9%. As before with three choices, classifications by random chance alone would be correct 33% of the time. Measurements from the zone corresponding to first-year lacustrine growth $\{FOUR(1)\}$ represented the most important variable in the models.

Scale Selection

The manner in which scales were selected for measurement had no significant effect on the ability of models to separate test scales according to their origins (Table 5). The mean classification accuracy was highest (93.0% to 97.3%) when Rules 2-4 (Geographical, Escapement, and Tag Probability Rules) were used to select scales from Canada for the NOM and was lowest (76.8 to 78.5%) when Rule 1 (Equal Probability Rule) was used to select Canadian scales for the SM. Classification accuracy varied as little as 1.9% to as much as 3.7% for different rules to select scales from Alaska waters; among Rules 2-4 applied to Canadian scales, the classification accuracy varied from 0.7 to 2.7%. When confidence intervals were placed around fractions, only 2 of the 48 90% confidence intervals (3 replications of 16 combinations) did not bracket the true fractions in the test population, which indicates that computed intervals correspond to about 95% confidence, not 90% (Table 6). With 90% confidence intervals, about five of the estimates are expected as outliers due to random chance alone¹. The two missed fractions occurred in the 90/10 splits when Rule 2 was applied to Alaskan scales and Rules 2 and 3 applied to scales from Canada.

The models that distinguish fish from Alaskan rivers, the Stikine River, and the Nass and Skeena Rivers as a unit (Rule 1, Canada scales) had negative fractions in 11 of 12 trials. Since all fractions must be between zero and unity, negative fractions indicated that no or few fish from the Stikine River were in the test population, which was the case here. When negative proportions in **u** occur in a model with K groups, the common practice is to remove the group with the negative proportion from the model and generate a new model on K-1 remaining groups. Application of Rule 2 to scales from Canada makes this adjustment. Cook (1983) developed an alternative procedure that adjusts the positive fractions in **u** for the negative fraction without building a new model.

DISCUSSION

Differences in Scale Patterns

Although only scales collected in 1982 were used in our analyses, our results indicate that differences in scale patterns are persistent from year to year. Because three year classes are represented in the age-pooled models, annual variation in scale patterns will reduce their accuracy over that achieved with

The conservative nature of confidence intervals around the proportions was not unexpected. Cook (1982) showed through Monte Carlo simulation that these confidence intervals, if calculated as suggested in Pella and Robertson (1979), are closer to 95% confidence intervals

Table 5. Comparison of mean accuracies (%) attained from models constructed from scales selected from stocks of sockeye salmon from both Alaska and from Canada according to four rules. The mean accuracies correspond to the comparisons made in Table 4, but unlike in Table 4 where accuracies are the result of classifying sets of scales with known origins, the mean accuracies here are the result of jackknife procedures used to build models.

			Ala		Donge	
	Rule	1	2	3	4	Range (Rules 1-4)
Canada	1	77.3	76.8	78.5	78.5	1.9
Canada	2	94.3	94.0	94.8	97.3	3.3
Canada	3	96.0	93.3	95.8	97.0	3.7
Canada	4	93.0	93.3	94.8	96.3	3.3
Range Rules	(2-4)	2.7	0.7	1.0	1.0	

Table 6. Comparison of four rules for selecting scales from stocks of sockeye salmon from Alaska and Canada against sets of scales with known origins. Each rule was applied to Alaska and to Canadian stocks and tested against three sets of scales (90/10, 50/50, and 10/90% Alaska/Canada). The intervals around the fractions are 90% confidence intervals. See text for description of rules and for description of tests.

	& AK/CA	Alaska Rule 1		Alaska Rule 1 Alaska Rule 2		Alaska	Rule 3	Alaska Rul	Alaska Rule 4		
Canada Rule l	90/10 50/50 10/90	Alaska Skeena/Nass .842 ± .150		1.104 ± .144 .01 .583 ± .148 .60	ena/Nasa Stikine 4 ± .108154 ± .147 19 ± .165192 ± .142 6 ± .167194 ± .177	.953 ± .131 .508 ± .139 .	keena/Nass Stikine 070 ± .105024 ± .122 555 ± .161063 ± .129 060 ± .152153 ± .149	.969 ± .115 .087	na/Nass Stikine ± .098056 ± .107 ± .159103 ± .127 ± .158170 ± .164		
Canada Rule 2	90/10 50/50 10/90	Alaska Skeena .854 ± .079 .146 ± .449 ± .054 .551 ± .056 ± .058 .944 ±	.079 .094	Alaska .977 ± .072 ,545 ± .096 .102 ± .061	.455 🛨 .096	Alaska .944 ± .0 .519 ± .0 .095 ± .0	67 .056 <u>+</u> .067 94 .481 <u>+</u> .094	Alaska .905 ± .053 .503 ± .088 .111 ± .066	Skeena/Nass .095 土 .053 .497 土 .088 .889 土 .066		
Canada Rule 3	90/10 50/50 10/90	Alaska Skeena .864 ± .075 .136 ± .473 ± .091 .527 ± .082 ± .053 .918 ±	.075 .091	Alaska 1.017 ± .075 .578 ± .099 .139 ± .065	.422 + .099	Alaska .940 <u>+</u> .0 .514 <u>+</u> .0 .109 <u>+</u> .0	63 .060 ± .063 92 .486 ± .092	Alaska .910 ± .054 .495 ± .089 .091 ± .063	Skeena/Nass .090 ± .054 .505 ± .089 .909 ± .063		
Canada Rule 4	90/10 50/50 10/90	Alaska Skeena .877 ± .086 .123 ± .471 ± .098 .529 ± .064 ± .060 .936 ±	086 098	Alaska .965 ± .075 .514 ± .098 .087 ± .064	.486 ± .098	Alaska .949 ± .0 .514 ± .0 .078 ± .0	65 .051 ± .065 94 .486 ± .094	Alaska .924 ± .057 .503 ± .090 .081 ± .063	Skeena/Nass .076 ± .057 .497 ± .090 .919 ± .063		

¹ Actual stock composition outside confidence interval boundaries.

age-specific models. The decrease in mean accuracy from the age-specific to the age-pooled models is, therefore, a measure of the importance of the variation in growth caused by density-independent (climate-caused) and density-dependent factors for the 1976 through 1978 year classes. The asymptotic mean accuracy dropped slightly (only 6% on average) between the age-specific and the age-pooled NOM and SM, which indicates that at least for 1976 through 1978, geographically related differences are far more important that annual variations in density-dependent and other density-independent factors.

Topography is probably the cause of the persistent differences in scale patterns. Rivers in Southeastern Alaska have a maritime climate with cool, wet, overcast summers and wet, overcast, mild winters. The rearing lakes of the Nass, Skeena, and Stikine Rivers have more continental weather with dry, warm, and sunny summers and cold, clear winters. These climatic differences most likely affect water temperatures, lacustrine productivity, length and intensity of growing seasons, and ultimately the growth of sockeye salmon during their freshwater years. Also, the long distance between rearing lakes in Canada and the sea (Figure 8) represents a growth environment for migrating sockeye salmon smolts that is not present in the shorter Alaskan rivers and is the most likely cause of the presence of plus growth in fish reared in Canada and the lack of plus growth in salmon reared in Southeastern Alaska.

Not only are geographic differences in scale patterns persistent between nations, but geographic similarities in scale patterns are persistent within nations. Preliminary comparisons of scale differences through one linear discriminant model showed only a few scales from a few Alaska rivers were misclassified as being from Canada. The comprehensive comparison of model accuracy, both mean accuracy and accuracy against an independent set of scales with known origins, showed that scales can be sampled in about any fashion with little effect on accuracy, even with the two most troublesome runs (Filmore and Helm Lake) represented in the comparison. For scales taken in 1982, any of the four rules of selection gave about the same results. In selecting scales from Canada, the major differences in the accuracies for the Equal Probability Rule versus the other rules are actually differences between the SM and the NOM. The insensitivity of model accuracy to the way scales are selected is a big advantage, because it permits the selection of scales by the most convenient way without a penalty in accuracy.

The persistence in scale patterns over year classes indicates that a historical model can be used to accurately estimate interception rates of sockeye salmon during a fishing season in southern Southeastern Alaska and northern British Columbia. This is fortunate because in many other fisheries, annual variation in scale patterns is too great to produce precise historical models while escapement samples are available too late to be of benefit for calculating interception rates within a season. Because differences and similarities in scale patterns are persistent across stocks from year to year (at least from 1976 through 1978), age-pooled models can be used to classify age-pooled samples from catches in southern Southeastern Alaska and northern British Columbia. After the season, age-specific NOM and SM can be built to remove annual variation in scale patterns from the estimates of interception rates and gain the added accuracy (in this study an average 6%). Because this study has analyzed only I year of data, only one age-pooled model is available for in-season work. As more information from future years becomes available, more historical comprehensive age-pooled models can be developed to estimate interception rates during seasons.

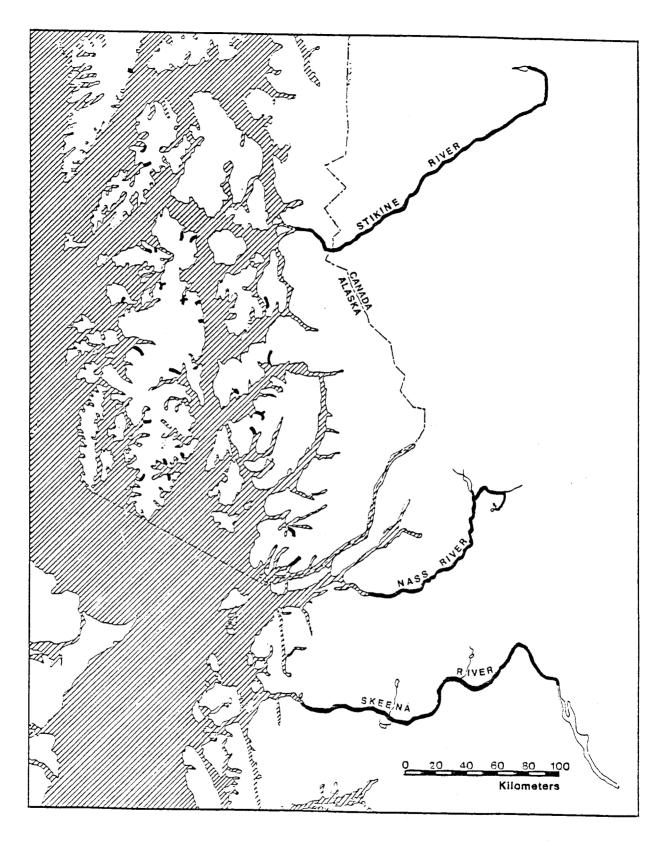


Figure 8. Comparative lengths of rivers in southern Southeastern Alaska and northern British Columbia that have major populations of sockeye salmon.

Other Stocks

In 1982, the stocks included in our analysis were those that tagging studies conducted that year indicated were present. However, preliminary results from the 1983 tagging program showed that significant numbers of sockeye salmon from the Fraser River in southern British Columbia were present in Alaska District 104. Therefore, scales from the escapement to the Fraser must be added to the Canada group to build a NOM or to the Nass/Skeena group to build a SM for 1983. Although no analysis has yet been done on the accuracy of models in separating these new combinations from Alaska scales, the prognosis is good. The Fraser, like the Nass, Stikine, and the Skeena, is a large, long river with rearing areas far from the ocean where climates are more continental. Based on this topography, the chances are good that growth and subsequently scale patterns on sockeye salmon reared in the Fraser River will be similar to patterns on scales from the Stikine, Nass, or Skeena Rivers.

Appropriate Models

The accuracy of all SM and NOM models is sufficient to provide highly precise estimates of interception rates for sockeye salmon caught in 1982. On average, models based on only three variables achieved mean accuracies of 91% and 78% for the age-specific NOM and SM, respectively. The average, asymptotic mean accuracies for the age-specific NOM and SM are 92 and 85%, respectively. The ability of SM and NOM to correctly classify sets of scales of known origin in 46 of 48 instances also shows that all the models are good estimator of interception rates.

Yet some models are more accurate than others. The two-choice National Origin Models were consistently more accurate than the three-choice Stikine Models. The SM models lose accuracy because scales from the Stikine and from the Nass and Skeena Rivers are very similar. The SM correctly classified scales from Alaskan rivers, but had some trouble with scales from the Canadian rivers. Since the topography, size, and climate of the Stikine River is very similar to that of the Nass and Skeena Rivers, there is little surprise that their growth and scale patterns were similar for 1982. However, this similarity is not so great as to prevent good discrimination of scale patterns. In eleven out of the twelve cases that the Equal Probability Rule was applied to scales from Canada, the SM indicated that no scales from the Stikine River were present when indeed they were not. In these cases, the correction matrix was good enough to compensate for the lower accuracies of the SM.

Because NOM have consistently better accuracy than SM, the latter models should be used only when fish from the Stikine River are in the district and there is a need to separate fish from this river from all others. Tagging studies show that sockeye salmon from the Stikine River were rare or not in Alaska Districts 101-4, 107, and Canadian Districts 1, 3, 4, and northern subdistricts of District 31 in 1982; the NOM models can provide highly accurate estimates of interception rates of fish from Alaska and the other two Canadian Rivers in these districts. Also, if NOM are used in these Districts, any strays from the Stikine River in these districts will most probably be classified as being from Canada and not from Alaska. In other fishing districts (Alaska Districts 105-6 and 108), accounting for sockeye salmon from the Stikine River is an important issue and SM should be used.

The incorrect classifications from the SM will be mostly among fish migrating to Canada.

RECOMMENDATIONS

- 1. Further Analysis of 1982 Data: Scale samples collected from the commercial catches in 1982 should be used to estimate the contribution of each nation's stocks to the fisheries of the other nation. Estimates should be made for each gear type and district on as fine a temporal scale as necessary.
- 2. Analysis of 1983 Data: Age-class specific models should be developed for the 1983 as was done for 1982 so that sources of variability in scale patterns among stocks, age classes, and year classes can be identified and their importance assessed. A new NOM will have to be developed for use in outer coastal fisheries which includes stocks from the Fraser River.
- 3. Sampling Program, 1984 (and beyond): Scale samples should be collected from principal escapements to provide data for age-specific models. After several seasons, sampling might possibly be reduced if interannual differences in scale patterns remains the same. Until then, commercial catch sampling programs should be conducted to provide weekly estimates of age composition in fisheries of concern to improve the precision of age-specific models. The temporal and spatial variability of interception rates within the principal fishing districts should be investigated with a comprehensive, sitespecific sampling program in 1984 for post-season analysis.
- 4. In-season Program, 1984: A program to estimate interception rates during a fishing season ought to be initiated during 1984. Target fisheries during the first year should include those in Alaska Districts 101, 104, and 106. The appropriate NOM or SM to use in-season will be determined following analysis of 1983 data. A comparison of estimates made during the conduct of the fisheries using historic models with estimates based on data from 1984 should be made.

ACKNOWLEDGMENTS

We are grateful to personnel of the Port Sampling Project of the Division of Commercial Fisheries, Southeastern Region, Alaska Department of Fish and Game; to personnel of the Fisheries Rehabilitation, Enhancement, and Development Division of the Alaska Department of Fish and Game; to employees of the Southern Southeast Regional Aquaculture Association; to the staff of the National Marine Fisheries Service, Auke Bay Laboratory; and to personnel of the Canadian Department of Fisheries and Oceans for their assistance in gathering samples from catches and escapements. Also, we wish to thank Mike Cable, Jan Weller, and Mike Hirn for their help in obtaining samples from escapements. For their assistance with data analysis, we gratefully acknowledge the help of Doug McBride, Ben Van Alen, and Samuel Sharr. And for their help with the manuscript, we thank Robert Conrad and Virginia Burton.

LITERATURE CITED

- Bilton, H. 1979. Data and computations used in racial analysis of age 1.2 and 1.3 North American sockeye salmon on the basis of their scale characters. Fish. Res. Bd. Can. Tech. Rept. No. 223, 57 pp.
- Cook, R. 1982. Estimating the mixing proportion of salmonids with scale pattern recognition applied to sockeye salmon (Oncorhynchus nerka) in and around the Japanese landbased driftnet fishery. Ph.D. Dissertation, Univ. of Washington, Seattle. 264 pp.
- Cook, R. 1983. Simulation and application of stock composition estimates. Cont. No. 626. Fish. Res. Inst., Univ. Washington, Seattle. 28 pp.
- Cook, R. and G. Lord. 1978. Identification of stocks of Bristol Bay sockeye salmon by evaluating scale patterns with a polynomial discriminant method. U.S. Fish Wildl. Serv., Fish. Bull. 76:415-423.
- Clutter, R. and L. Whitesel. 1956. Collection and interpretation of sockeye salmon scales. Bull. Int. Pac. Salmon Fish. Comm., No. 9. 159 pp.
- Dixon, W. and M. Brown. 1979. Biomedical computer programs p-series. Univ. of Calif. Press, Berkeley. 880 pp.
- Fisher, R. 1936. The use of multiple measurements in taxonomic problems. Animal Eugenics 7:179-188.
- Lachenbruch, P. 1975. Discriminant analysis. Hafner Press, New York. 128 pp.
- Logan, R. 1967. The Noyes Island salmon fishing conflict. British Columbia Geographic Series, Occasional Paper 8:18-26.
- McGregor, A. 1983. Age, sex, and size of sockeye salmon (Oncorhynchus nerka Walbaum) catches and escapements in Southeastern Alaska in 1982. Alaska Dept. of Fish and Game, Tech. Data Rpt. No. 100. 124 pp.
- Mosher, K. 1968. Photographic atlas of sockeye salmon scales. U.S. Fish and Wildl. Serv., Fish. Bull. 67:243-280.
- Noerenberg, W. 1959. Salmon migrations in Southeastern Alaska. U.S. Fish and Wildl. Circ. No. 59.
- Oliver, G., S. McPherson, and S. Marshall. 1983. Progress Report: Feasibility of determining the origin of sockeye salmon (Oncorhynchus nerka) in the southern Southeast Alaska Districts 101-108 net fisheries using scale pattern analysis. 1982. Alaska Dept. of Fish and Game, Juneau. Unpublished Manuscript. 92 pp.
- Pella, J. and T. Robertson. 1979. Assessment of composition of stock mixtures. Fishery Bull. 77:387-389.

- Rich, W. and F. Morton. 1930. Salmon-tagging experiments in Alaska, 1927-1928. U.S. Bur. Fish. Bull. 45:1-23.
- Ryan, P. and M. Christie. 1976. Scale reading equipment. Fish and Mar. Serv., Canada, Tech. Rept. No. PAC/T-75-8. 38 pp.
- Simpson, L. 1968. Sockeye salmon migratory behavior and biological statistics collection, Southeastern Alaska. Ann. Prog. Rept. for period July 1, 1967 to June 30, 1968. Anadromous Fish Act (P.L. 89-304), Proj. No. AFC-2-2. 6 pp.
- Tesch, F. 1970. Age and growth, p. 98-126. <u>In Methods for assessment of fish productions in fresh waters</u>, W.E. Richer (ed.) IBP Handbook No. 3, Blackwell Scientific Publications, Oxford and Edinburgh.
- Verhoeven, L. 1952. A report to the salmon fishing industry of Alaska on the results of the 1947 tagging experiments. Unpub. Manuscript., Univ. of Wash. Fish. Res. Inst., Seattle.

The Alaska Department of Fish and Game administers all programs and activities free from discrimination based on race, color, national origin, age, sex, religion, marital status, pregnancy, parenthood, or disability. The department administers all programs and activities in compliance with Title VI of the Civil Rights Act of 1964, Section 504 of the Rehabilitation Act of 1973, Title II of the Americans with Disabilities Act of 1990, the Age Discrimination Act of 1975, and Title IX of the Education Amendments of 1972.

If you believe you have been discriminated against in any program, activity, or facility, or if you desire further information please write to ADF&G, P.O. Box 25526, Juneau, AK 99802-5526; U.S. Fish and Wildlife Service, 4040 N. Fairfax Drive, Suite 300 Webb, Arlington, VA 22203 or O.E.O., U.S. Department of the Interior, Washington DC 20240.

For information on alternative formats for this and other department publications, please contact the department ADA Coordinator at (voice) 907-465-6077, (TDD) 907-465-3646, or (FAX) 907-465-6078.